SIAM J. SCI. COMPUT.
Vol. 35, No. 2, pp. A785–A814

ANALYSIS OF THE MINIMAL RESIDUAL METHOD APPLIED TO ILL POSED OPTIMALITY SYSTEMS*

BJØRN FREDRIK NIELSEN† AND KENT-ANDRE MARDAL‡

Abstract. We analyze the performance of the minimal residual (MINRES) method applied to linear Karush–Kuhn–Tucker systems arising in connection with inverse problems. Such optimality systems typically have a saddle point structure and have unique solutions for all $\alpha > 0$, where α is the parameter employed in the Tikhonov regularization. Unfortunately, the associated spectral condition number is very large for small values of α, which strongly indicates that their numerical treatment is difficult. Our main result shows that a broad range of linear ill posed optimality systems can be solved efficiently with the MINRES method. This result is obtained by carefully analyzing the spectrum of the associated saddle point operator: Except for a few isolated eigenvalues, the spectrum consists of three bounded intervals. Krylov subspace methods handle such problems very well. For severely ill posed cases, techniques based on Chebyshev polynomials are applied to prove that the number of iterations needed by the MINRES method cannot grow faster than $O([\ln(\alpha^{-1})]^2)$ as $\alpha \to 0$. We illuminate our analysis with numerical results for inverse problems involving partial differential equations. In these examples we observe that the required number of iterations is almost of order $O(\ln(\alpha^{-1}))$ for moderately sized α, and we discuss how this behavior is linked to our theoretical findings.

Key words. PDE constrained optimization, Karush–Kuhn–Tucker (KKT) systems, Krylov subspace methods, inverse problems, all-at-once method

AMS subject classifications. 49K20, 65F08, 65N21, 65J22, 65F15

DOI. 10.1137/120871547

1. Introduction.

In recent years many researchers have studied the numerical treatment of inverse problems, especially the minimization of quadratic cost functionals with constraints expressed in terms of PDEs—so-called PDE constrained optimization [8, 9, 15, 20, 40]. This type of problem can be solved with iterative minimization methods, or one can use the Lagrange multiplier technique to obtain a system of equations which must be satisfied by the optimal solution. If the latter approach is applied, discretization leads to a large system of algebraic equations, and the optimization problem can be solved with an all-at-once method, that is, a method in which the optimality condition, the state equation, and its adjoint, which constitute the optimality system, are solved in a fully implicit manner.

The optimality system inherits the ill posed nature of the underlying inverse problem, and regularization techniques must therefore be invoked. If Tikhonov regularization is applied, then the spectral condition number of the system typically is of order $O(\alpha^{-1})$, where $\alpha > 0$ denotes the regularization parameter. The purpose of adding the regularization term is to obtain a well posed problem with a reasonable condition number. In practice one would therefore not choose α too small, but $\alpha \in (10^{-4}, 10^{-2})$ is not unlikely,[1] which often will lead to a relatively large condition

*Submitted to the journal's Methods and Algorithms for Scientific Computing section March 27, 2012; accepted for publication (in revised form) December 17, 2012; published electronically March 19, 2013.

http://www.siam.org/journals/sisc/35-2/87154.html

†Department of Mathematical Sciences and Technology, Norwegian University of Life Sciences, 1432 Aas, Norway; Simula Research Laboratory, 1325 Lysaker, Norway; and Center for Cardiological Innovation, Oslo University Hospital, Oslo, Norway (bjorn.f.nielsen@umb.no).

‡Center for Biomedical Computing, Simula Research Laboratory, 1325 Lysaker, Norway; and Department of Informatics, University of Oslo, 0316 Oslo, Norway (kent-and@simula.no).

[1]The appropriate size of the regularization parameter depends on the amount of noise in the input data [17].

number for the optimality system. Furthermore, according to the standard theory for Krylov subspace methods, the number of iterations required by the minimal residual (MINRES) method grows rapidly as the condition number increases. This suggests that the MINRES scheme might not be well suited for solving regularized Karush–Kuhn–Tucker (KKT) problems arising in engineering. In this text we will show that this rough analysis is not accurate, and that it provides an unrealistic overestimate for the workload needed by the MINRES algorithm.

The purpose of this paper is to analyze the MINRES method [28] applied to a large class of ill posed linear optimality systems. Let

- H_1 be the parameter/control space,
- H_2 the state space, and
- H_3 the observation space,

with norms $\|\cdot\|_{H_1}$, $\|\cdot\|_{H_2}$, and $\|\cdot\|_{H_3}$. We study the numerical treatment of

$$\min_{v\in H_1,\, u\in H_2} \left\{ \frac{1}{2}\|Tu-d\|_{H_3}^2 + \frac{1}{2}\alpha\|v\|_{H_1}^2 \right\} \tag{1.1}$$

subject to

$$Au = -Bv \quad \text{(state equation)}. \tag{1.2}$$

The quantity d is given and $\alpha \geq 0$ is a regularization parameter. These types of problems typically arise when one wants to use an observation d to recover the parameter/control v in the state equation.

If

$$\mathcal{A}_\alpha p = b \tag{1.3}$$

is the optimality system associated with (1.1)–(1.2), then the main objectives of this text can be roughly formulated as follows:

(a) We prove that the spectrum of $\mathcal{A}_\alpha$ is *almost* contained in bounded intervals:

$$\mathrm{sp}(\mathcal{A}_\alpha) \subset [-b,-a] \cup [c\alpha, d\alpha] \cup \{\lambda_1, \lambda_2, \ldots, \lambda_{N(\alpha)}\} \cup [a,b],$$

where $N(\alpha)$ is of order $O(\ln(\alpha^{-1}))$ in the severely ill posed case and a, b, c, $d > 0$ are constants that do not depend on α. Krylov subspace solvers are well known to handle problems with few isolated eigenvalues excellently; see e.g., [4]. This matter is discussed in detail for the MINRES method in this paper. We have earlier studied a particular preconditioning strategy for ill posed KKT systems that leads to (theoretical) iteration counts of order $O([\ln(\alpha)]^2)$; see [27]. The present text may be regarded as a follow-up paper to that article.

(b) Through numerical examples we show how (a) can be employed to solve PDE constrained optimization problems efficiently.

Our investigation is presented in terms of functional analysis and uses basic properties of inverse problems. The results are therefore applicable whenever a problem can be written in the form (1.1)–(1.2). Nevertheless, we were motivated by practical experience with PDE constrained optimization. For such problems, not only do issues arise for small $\alpha > 0$, but the condition number of the associated KKT system will typically increase significantly as the mesh parameter $h > 0$, used in the discretization of the involved PDE, decreases. For some model problems, we explain how this latter matter can be handled by invoking multigrid preconditioners.

Note that our analysis addresses only Tikhonov regularization. As far as the authors know, it is not straightforward to generalize our findings to cover other regularization methods. Separate investigations devoted to other techniques are therefore needed. Nevertheless, since our methodology mainly employs basic properties of inverse problems and self-adjoint operators, it is likely that similar results can be established for a variety of regularization methods.

The numerical treatment of saddle point operators arising in connection with PDEs is a contemporary research field [2, 6, 11, 13, 14, 18, 31, 34, 42]; see [24] for a rather recent review. For well posed problems one of the main issues is to obtain iteration counts that are acceptable as the mesh parameter $h > 0$ decreases. If the parameter identification task at hand is ill posed, then one must also ensure that the iterative schemes can handle cases with small regularization parameters, i.e., that the number of iterations needed does not increase significantly as $\alpha \to 0$. The latter type of problem has been addressed in many papers for various models [1, 7, 19, 25, 29, 30, 32, 33, 35], i.e., for special cases of elliptic and parabolic control problems. For some particular cost functionals, remarkable algorithms that are completely robust with respect to α have been developed [32, 33, 36, 43]. Also the spectral properties of $\mathcal{A}_\alpha$ have been analyzed thoroughly for the case in which (1.2) is the Poisson problem; see Thorne [37, 38]. In this paper an abstract approach is used to cover a rather broad range of saddle point problems, and we conclude that Krylov subspace solvers might be an attractive alternative for their effective numerical solution.

If one wants to solve a practical problem involving real world data, it is almost certainly not sufficient to solve (1.3) once with one particular choice of α. In fact, procedures for estimating an appropriate size of the regularization parameter typically requires that (1.3) be solved repeatedly for a sequence of different values of α; see e.g., [17]. We may thus conclude that the efficiency needed to solve an inverse problem is of a different magnitude from what is required for a well posed problem—the fast numerical solution of (1.3) is crucial.

The next section contains all but one of the assumptions that we need. Notation for the optimality system is introduced in section 3, which also contains the final assumption. The eigenvalue distribution of the indefinite optimality system (1.3) is analyzed in section 4, and our numerical experiments are presented in section 5. Finally, section 6 is devoted to the theoretical convergence behavior of the MINRES method.

2. Assumptions. Throughout this text c, $\tilde{c}$, C, and $\tilde{C}$ are (generic) positive constants that do not depend on the regularization parameter α. We limit our analysis to linear state equations (1.2) and assume that

($\mathcal{A}1$) $A : H_2 \to H_2$ is bounded and linear,[2]
($\mathcal{A}2$) A^{-1} exists and is bounded,
($\mathcal{A}3$) $B : H_1 \to H_2$ is bounded and linear,
($\mathcal{A}4$) $T : H_2 \to H_3$ is bounded and linear, and
($\mathcal{A}5$) the inf-sup condition holds:

$$\inf_{w \in H_2} \sup_{(v,u) \in H_1 \times H_2} \frac{(Bv, w) + (Au, w)}{\sqrt{\|v\|_{H_1}^2 + \|u\|_{H_2}^2}\, \|w\|_{H_2}} \geq c. \tag{2.1}$$

[2]If the state equation (1.2) is a PDE, then A will typically be a mapping from H_2 to its dual space H_2'. An operator $R_2^{-1} : H_2' \to H_2$ must thus be applied to the state equation in order to get a mapping $R_2^{-1}A : H_2 \to H_2$. One may regard R_2^{-1} to be a preconditioner [24]. In the present text, R_2^{-1} will be the inverse of the Riesz map or a suitable multigrid approximation of this operator. We will return to this issue in the numerical experiment section.

Note that A is the mapping that must be inverted in order to solve the state equation, B maps the (unknown) parameter/control v into the state equation, and T is the observation operator.

From $\mathcal{A}2$ and $\mathcal{A}3$ it follows that the solution u of (1.2) depends continuously on v:

$$\|u\|_{H_2} \leq C\|v\|_{H_1}, \tag{2.2}$$

i.e., the state equation is well posed.

3. Optimality system and one more assumption. The optimality system associated with (1.1)–(1.2) can be derived by employing standard techniques. Details about this issue can be found in, e.g., [39]. Here we merely state the result. That is, the solution of (1.1)–(1.2) must solve the following problem: Find $(v, u, w) \in H_1 \times H_2 \times H_2$ such that

$$\begin{bmatrix} \alpha I & 0 & B^* \\ 0 & T^*T & A^* \\ B & A & 0 \end{bmatrix} \begin{bmatrix} v \\ u \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ T^*d \\ 0 \end{bmatrix}, \tag{3.1}$$

where w is the Lagrange multiplier and the "*" notation denotes adjoint. It is well known that this system typically is indefinite and hence has both positive and negative eigenvalues.

Remark. Assume that (1.2) is a PDE. Then boundary conditions would typically lead to a nonzero contribution to the third position of the right-hand-side vector of (3.1). For the analysis presented in this paper, this is of no importance because we use spectral properties to study the convergence of the MINRES method.

For the sake of convenience, we introduce the notation

$$\mathcal{A}_\alpha = \begin{bmatrix} \alpha I & 0 & B^* \\ 0 & T^*T & A^* \\ B & A & 0 \end{bmatrix} \quad \text{for } \alpha \geq 0, \tag{3.2}$$

$$p = \begin{bmatrix} v \\ u \\ w \end{bmatrix},$$

$$b = \begin{bmatrix} 0 \\ T^*d \\ 0 \end{bmatrix},$$

and (3.1) can be written in the form

$$\mathcal{A}_\alpha p = b. \tag{3.3}$$

Note that

$$\mathcal{A}_\alpha : (H_1 \times H_2 \times H_2) \to (H_1 \times H_2 \times H_2),$$

and that for $\alpha = 0$ we get

$$\mathcal{A}_0 = \begin{bmatrix} 0 & 0 & B^* \\ 0 & T^*T & A^* \\ B & A & 0 \end{bmatrix}, \tag{3.4}$$

which contains zero regularization.

Throughout this text we consider problems of the form (1.1)–(1.2) that are ill posed for $\alpha = 0$. This undesirable property is likely to be inherited by the optimality system (3.1):

($\mathcal{A}6$) We assume that $\mathcal{A}_0$ is a compact operator, i.e., $\lambda_i(\mathcal{A}_0) \to 0$ as $i \to \infty$. Furthermore, in the severely ill posed case the eigenvalues are assumed to satisfy

$$|\lambda_i(\mathcal{A}_0)| \le c\, e^{-Ci} \quad \text{for } i = 1, 2, \dots. \tag{3.5}$$

Here, c and C are positive constants not depending on α, since $\mathcal{A}_0$ does not involve α.

We also assume the spectrum of $\mathcal{A}_\alpha$ is discrete/countable for every $\alpha > 0$.

If, unexpectedly, the ill-posedness of (1.1)–(1.2) for $\alpha = 0$ is not inherited by (3.1), then the numerical solution of the latter is of course much easier.

In a more general version of $\mathcal{A}6$ one could assume that the spectrum of $\mathcal{A}_0$ contains only a subsequence that fulfills a bound of the form (3.5). This, however, leads to an even more involved analysis which would have to address noncompact operators. (Moreover, for problems with continuous spectra it is not even straightforward to define the concepts *severely* and *mildly* ill posed [21].)

4. Eigenvalues. In this section we characterize the basic structure of the spectrum of $\mathcal{A}_\alpha$, defined in (3.2). This information will be used in section 6 to analyze the convergence behavior of the MINRES method applied to the saddle point problem (3.1), or, equivalently, applied to (3.3).

4.1. Basic bounds. In Appendix A standard techniques for saddle point operators are used to obtain bounds for the operator norms of $\mathcal{A}_\alpha$ and $\mathcal{A}_\alpha^{-1}$:

$$\begin{aligned} &\|\mathcal{A}_\alpha\| \le C \quad \text{for all } \alpha \in [0, 1], \\ &\|\mathcal{A}_\alpha^{-1}\| \le \frac{1}{c\alpha} \quad \text{for all } \alpha \in (0, 1]. \end{aligned} \tag{4.1}$$

We will now employ this information to analyze the eigenvalues of $\mathcal{A}_\alpha$.

If

$$\mathcal{A}_\alpha q = \lambda q,$$

then

$$|\lambda| \|q\| = \|\mathcal{A}_\alpha q\|,$$

or

$$|\lambda| = \frac{\|\mathcal{A}_\alpha q\|}{\|q\|} \le \frac{\|\mathcal{A}_\alpha\| \|q\|}{\|q\|} = \|\mathcal{A}_\alpha\| \le C.$$

Likewise,

$$\mathcal{A}_\alpha q = \lambda q$$

implies that

$$\frac{1}{\lambda} q = \mathcal{A}_\alpha^{-1} q.$$

That is

$$\frac{1}{|\lambda|} \|q\| \le \|\mathcal{A}_\alpha^{-1}\| \|q\|,$$

or

$$|\lambda| \geq \frac{1}{\|\mathcal{A}_\alpha^{-1}\|} \geq c\alpha.$$

LEMMA 4.1. *Let $\mathcal{A}_\alpha$ be the operator defined in* (3.2). *There exist constants c, $C > 0$, which are independent of $\alpha \in [0,1]$, such that*

$$c\alpha \leq |\lambda_i(\mathcal{A}_\alpha)| \leq C \quad \textit{for } i = 1, 2, \ldots.$$

Note that these bounds also hold for $\alpha = 0$.

The numerical values of c and C depend on the problem under consideration. For example, if the state equation is a PDE, then the variation of the involved coefficient functions and the geometrical properties will have significant impact on the size of these constants. Section 5 contains information about this issue for some test problems; see Figures 2, 4, and 6.

4.2. Negative eigenvalues. The next step is to prove that the negative eigenvalues of $\mathcal{A}_\alpha$ are well behaved for all $\alpha \geq 0$. More specifically, we will show that the negative eigenvalues cannot approach zero as $\alpha \to 0$. Our analysis employs the following auxiliary result.

LEMMA 4.2. *Let A be the operator in the state equation* (1.2). *Assumptions $\mathcal{A}1$ and $\mathcal{A}2$ imply that AA^* is coercive, i.e., there exists a constant $c > 0$ such that*

$$(AA^*\phi, \phi) \geq c\|\phi\|_{H_2}^2 \quad \textit{for all } \phi \in H_2. \tag{4.2}$$

Proof. Assume that there does *not* exist a constant $c > 0$ such that (4.2) holds. Then there exists a sequence $\{\phi_n\}_{n=1}^\infty$, with $\|\phi_n\|_{H_2} = 1$, such that

$$(AA^*\phi_n, \phi_n) \longrightarrow 0 \quad \text{as } n \to \infty.$$

Let us consider

$$\frac{\|(A^*)^{-1}y\|_{H_2}^2}{\|y\|_{H_2}^2}$$

with $y = y_n = A^*\phi_n$:

$$\begin{aligned}
\frac{\|(A^*)^{-1}y_n\|_{H_2}^2}{\|y_n\|_{H_2}^2} &= \frac{((A^*)^{-1}y_n, (A^*)^{-1}y_n)}{(y_n, y_n)} \\
&= \frac{(\phi_n, \phi_n)}{(A^*\phi_n, A^*\phi_n)} \\
&= \frac{(\phi_n, \phi_n)}{(AA^*\phi_n, \phi_n)} \\
&= \frac{1}{(AA^*\phi_n, \phi_n)} \longrightarrow \infty \quad \text{as } n \to \infty.
\end{aligned}$$

Hence, $(A^*)^{-1}$ is not bounded, which contradicts assumption $\mathcal{A}2$. We conclude that AA^* must be coercive. ∎

(This lemma can also be established by using the bounded inverse theorem.)

The result regarding the negative eigenvalues of $\mathcal{A}_\alpha$ reads as follows.

LEMMA 4.3. *There exist constants $a, b > 0$ such that all the negative eigenvalues of $\mathcal{A}_\alpha$ are contained in the interval $[-b, -a]$. These constants do not depend on the size of the regularization parameter $\alpha \in [0, 1]$. (Note that $[-b, -a]$ also contains all the negative eigenvalues of $\mathcal{A}_0$.)*

Proof. Assume that $\lambda < 0$ is a negative eigenvalue of $\mathcal{A}_\alpha$ with associated eigenfunction $(v, u, w)^T$, i.e.,

$$\begin{bmatrix} \alpha I & 0 & B^* \\ 0 & T^*T & A^* \\ B & A & 0 \end{bmatrix} \begin{bmatrix} v \\ u \\ w \end{bmatrix} = \lambda \begin{bmatrix} v \\ u \\ w \end{bmatrix}$$

or

$$\begin{aligned} \alpha v + B^* w &= \lambda v, \\ T^*Tu + A^* w &= \lambda u, \\ Bv + Au &= \lambda w. \end{aligned}$$

Since $\lambda < 0$, $\lambda I - T^*T$ is invertible and it follows that

$$v = \frac{1}{\lambda - \alpha} B^* w, \tag{4.3}$$

$$u = (\lambda I - T^*T)^{-1} A^* w, \tag{4.4}$$

$$Bv + Au = \lambda w. \tag{4.5}$$

Note that $w = 0$ implies that $u = 0$ and $v = 0$, and we may assume that $w \neq 0$. By inserting expressions (4.3) and (4.4) for v and u, respectively, into (4.5) one finds that

$$\frac{1}{\lambda - \alpha} BB^* w + A(\lambda I - T^*T)^{-1} A^* w = \lambda w$$

or

$$-\lambda w = -\frac{1}{\lambda - \alpha} BB^* w + A(T^*T - \lambda I)^{-1} A^* w. \tag{4.6}$$

The next step is to discuss the properties of $T^*T - \lambda I$. Thereafter we return to (4.6). Recall that $\lambda < 0$ and therefore $T^*T - \lambda I$ is positive definite. Lemma 4.1 states that

$$|\lambda| \leq C. \tag{4.7}$$

We conclude that the spectrum $\mathrm{sp}(T^*T - \lambda I)$ of $T^*T - \lambda I$ satisfies

$$\mathrm{sp}(T^*T - \lambda I) \subset (0, \|T^*T\| + C] .$$

It follows that $(T^*T - \lambda I)^{-1}$ also is positive definite and that

$$\mathrm{sp}\left((T^*T - \lambda I)^{-1}\right) \subset \left[\frac{1}{\|T^*T\| + C}, \infty \right).$$

If we combine this information with (4.6), we find that

$$\begin{aligned}
-\lambda(w,w)_{H_2} &= -\frac{1}{\lambda-\alpha}(BB^*w,w)_{H_2} + (A(T^*T-\lambda I)^{-1}A^*w,w)_{H_2} \\
&= -\frac{1}{\lambda-\alpha}(B^*w,B^*w)_{H_1} + ((T^*T-\lambda I)^{-1}A^*w,A^*w)_{H_2} \\
&\geq ((T^*T-\lambda I)^{-1}A^*w,A^*w)_{H_2} \\
&\geq \frac{1}{\|T^*T\|+C}(A^*w,A^*w)_{H_2} \\
&= \frac{1}{\|T^*T\|+C}(AA^*w,w)_{H_2} \\
&\geq \frac{c}{\|T^*T\|+C}(w,w)_{H_2},
\end{aligned}$$

where we have used that $\lambda-\alpha<0$—recall that $\lambda<0$ and $\alpha\geq 0$—and inequality (4.2) in Lemma 4.2. Consequently,

$$\lambda \leq -\frac{c}{\|T^*T\|+C},$$

which together with (4.7) finishes the proof, i.e.,

$$\begin{aligned}
a &= \frac{c}{\|T^*T\|+C}, \\
b &= C. \qquad \square
\end{aligned}$$

4.3. Positive eigenvalues. The negative eigenvalues of $\mathcal{A}_\alpha$ are well behaved regardless of the size of the regularization parameter $\alpha\in[0,1]$. On the other hand, we have assumed that zero is a cluster point of the spectrum of $\mathcal{A}_0$; see assumption $\mathcal{A}6$. We will now investigate in what sense this property of $\mathcal{A}_0$ is inherited by the positive eigenvalues of $\mathcal{A}_\alpha$.

Note that

$$\begin{aligned}
\mathcal{A}_\alpha &= \begin{bmatrix} 0 & 0 & B^* \\ 0 & T^*T & A^* \\ B & A & 0 \end{bmatrix} + \begin{bmatrix} \alpha I & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\
&= \mathcal{A}_0 + \mathcal{E}_\alpha,
\end{aligned}$$

where

$$\mathcal{E}_\alpha = \begin{bmatrix} \alpha I & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

and we conclude that the difference between $\mathcal{A}_\alpha$ and $\mathcal{A}_0$ is small, provided that α is small. Is also the difference between the eigenvalues of these two operators small? Yes, indeed. As we will now briefly discuss, the min-max theorem (Courant–Fischer–Weyl min-max principle) [26, 41] provides a strategy for analyzing this issue. The conclusion is that the difference between the eigenvalues, properly sorted, of $\mathcal{A}_\alpha$ and $\mathcal{A}_0$ cannot be larger than α.

Even though we have assumed that $\mathcal{A}_0$ is compact, $\mathcal{A}_\alpha$ will in general not be compact for $\alpha>0$. In the case of infinite-dimensional spaces one can thus not (directly) apply the classical Courant–Fischer–Weyl min-max principle to $\mathcal{A}_\alpha$. For the sake of

simplicity, we will therefore now address only the finite-dimensional setting, which allows the use of the principle. (Because the min-max approach also is applicable to the discrete end of the spectrum of self-adjoint operators that are bounded below, similar results can be established in the infinite-dimensional case. More specifically, both $cI - \mathcal{A}_\alpha$ and $cI - \mathcal{A}_0$ are bounded below for a sufficiently large constant $c > 0$.)

Let

$$\lambda_1^+(\mathcal{A}_\alpha) \geq \lambda_2^+(\mathcal{A}_\alpha) \geq \cdots$$

and

$$\lambda_1^+(\mathcal{A}_0) \geq \lambda_2^+(\mathcal{A}_0) \geq \cdots$$

be the nonnegative eigenvalues of $\mathcal{A}_\alpha$ and $\mathcal{A}_0$, respectively, sorted in decreasing order. Note that $\mathcal{A}_\alpha$ and $\mathcal{A}_0$ are Hermitian. According to the min-max theorem (see, e.g., [41, 26]),

$$\lambda_i^+(\mathcal{A}_\alpha) = \max_{S_i} \min_{p \in S_i, \|p\|=1} (\mathcal{A}_\alpha p, p), \tag{4.8}$$

$$\lambda_i^+(\mathcal{A}_\alpha) = \min_{S_{i-1}} \max_{p \in S_{i-1}^\perp, \|p\|=1} (\mathcal{A}_\alpha p, p), \tag{4.9}$$

$$\lambda_i^+(\mathcal{A}_0) = \max_{S_i} \min_{p \in S_i, \|p\|=1} (\mathcal{A}_0 p, p), \tag{4.10}$$

$$\lambda_i^+(\mathcal{A}_0) = \min_{S_{i-1}} \max_{p \in S_{i-1}^\perp, \|p\|=1} (\mathcal{A}_0 p, p), \tag{4.11}$$

where S_i and S_{i-1} denote i-dimensional and $(i-1)$-dimensional subspaces of $H_1 \times H_2 \times H_2$, respectively. Now,

$$(\mathcal{A}_\alpha p, p) = (\mathcal{A}_0 p, p) + (\mathcal{E}_\alpha p, p),$$

which implies that

$$\begin{aligned}
\max_{p \in S_{i-1}^\perp, \|p\|=1} (\mathcal{A}_\alpha p, p) &= \max_{p \in S_{i-1}^\perp, \|p\|=1} \{(\mathcal{A}_0 p, p) + (\mathcal{E}_\alpha p, p)\} \\
&\leq \max_{p \in S_{i-1}^\perp, \|p\|=1} (\mathcal{A}_0 p, p) + \max_{p \in S_{i-1}^\perp, \|p\|=1} (\mathcal{E}_\alpha p, p) \\
&\leq \max_{p \in S_{i-1}^\perp, \|p\|=1} (\mathcal{A}_0 p, p) + \alpha \\
&\Downarrow \\
\min_{S_{i-1}} \max_{p \in S_{i-1}^\perp, \|p\|=1} (\mathcal{A}_\alpha p, p) &\leq \min_{S_{i-1}} \max_{p \in S_{i-1}^\perp, \|p\|=1} (\mathcal{A}_0 p, p) + \alpha \\
&\Downarrow \\
\lambda_i^+(\mathcal{A}_\alpha) &\leq \lambda_i^+(\mathcal{A}_0) + \alpha,
\end{aligned}$$

where we have used (4.9) and (4.11). In a similar fashion one can employ (4.8) and (4.10) to show that

$$\lambda_i^+(\mathcal{A}_\alpha) \geq \lambda_i^+(\mathcal{A}_0),$$

and we conclude that

$$0 \leq \lambda_i^+(\mathcal{A}_0) \leq \lambda_i^+(\mathcal{A}_\alpha) \leq \lambda_i^+(\mathcal{A}_0) + \alpha \quad \text{for } i = 1, 2, \ldots. \tag{4.12}$$

(Inequalities of this kind for Hermitian matrices are discussed on page 396 in Golub and Van Loan [16].)

Recall that we have assumed that the eigenvalues of $\mathcal{A}_0$ decay exponentially:

$$|\lambda_i(\mathcal{A}_0)| \le \tilde{c}\, e^{-\tilde{C}i} \quad \text{for } i = 1, 2, \ldots. \tag{4.13}$$

But $\{\lambda_i^+(\mathcal{A}_0)\}$ is a subsequence of $\{|\lambda_i(\mathcal{A}_0)|\}$ and therefore

$$\lambda_i^+(\mathcal{A}_0) \le \tilde{c}\, e^{-\tilde{C}i} \quad \text{for } i = 1, 2, \ldots.$$

If this bound is combined with inequalities (4.12) and Lemma 4.1, we obtain the following lemma.

LEMMA 4.4. *The nonnegative eigenvalues of $\mathcal{A}_\alpha$ and $\mathcal{A}_0$, sorted in decreasing order, satisfy*

$$c\alpha \le \lambda_i^+(\mathcal{A}_\alpha) \le \lambda_i^+(\mathcal{A}_0) + \alpha \le \tilde{c}\, e^{-\tilde{C}i} + \alpha \quad \textit{for } i = 1, 2, \ldots. \tag{4.14}$$

Remember that the negative eigenvalues of both $\mathcal{A}_\alpha$ and $\mathcal{A}_0$ are well behaved. Inequalities (4.14) therefore show that the eigenvalues of $\mathcal{A}_0$ that cluster at zero lead to eigenvalues of $\mathcal{A}_\alpha$ that are contained in an interval of the form $[c\alpha, d\alpha]$. For example, let $N = N(\alpha)$ be the smallest positive integer such that

$$\tilde{c}\, e^{-\tilde{C}N(\alpha)} \le \alpha;$$

then $\lambda_i^+(\mathcal{A}_\alpha) \in [c\alpha, 2\alpha]$ for all $i \ge N(\alpha)$. Moreover, $N(\alpha)$ is of order $O(\ln(\alpha^{-1}))$.

Remark. The Courant–Fischer–Weyl min-max principle can of course also be used to study the negative elements of the spectra of $\mathcal{A}_\alpha$ and $\mathcal{A}_0$. Since the analysis is similar to the one presented above, it suffices to state the following result:

$$\lambda_i^-(\mathcal{A}_0) \le \lambda_i^-(\mathcal{A}_\alpha) \le \lambda_i^-(\mathcal{A}_0) + \alpha \quad \text{for } i = 1, 2, \ldots. \tag{4.15}$$

Roughly speaking, (4.12) and (4.15) show that the difference between the eigenvalues of $\mathcal{A}_\alpha$ and $\mathcal{A}_0$ is of order $O(\alpha)$.

4.4. Spectrum. We now know that

- $\mathrm{sp}(\mathcal{A}_\alpha) \subset [-C, C]$ (see Lemma 4.1),
- the negative eigenvalues of $\mathcal{A}_\alpha$ are contained in $[-\tilde{b}, -\tilde{a}]$ (see Lemma 4.3),
- the positive eigenvalues of $\mathcal{A}_\alpha$ satisfy

$$c\alpha \le \lambda_i^+(\mathcal{A}_\alpha) \le \tilde{c}\, e^{-\tilde{C}i} + \alpha \quad \text{for } i = 1, 2, \ldots$$

  (see Lemma 4.4).

If we choose

$$b = \max\{\tilde{b}, C\}, \tag{4.16}$$

$$a = \tilde{a}, \tag{4.17}$$

we thus obtain our main result.

THEOREM 4.5. *The spectrum of $\mathcal{A}_\alpha$ satisfies*

$$\mathrm{sp}(\mathcal{A}_\alpha) \subset [-b, -a] \cup [c\alpha, 2\alpha] \cup \{\lambda_1, \lambda_2, \ldots, \lambda_{N(\alpha)}\} \cup [a, b], \tag{4.18}$$

where

$$2\alpha < \lambda_i < a \quad \textit{for } i = 1, 2, \ldots, N(\alpha),$$
$$N(\alpha) \le \left\lceil \frac{\ln(\tilde{c}) - \ln(\alpha)}{\tilde{C}} \right\rceil = O(\ln(\alpha^{-1})),$$

provided that $\alpha \in (0, 1]$. The constants a, b, c, $\tilde{c}$, $\tilde{C} > 0$ do not depend on α.

Remark. For the sake of convenience we choose a, which is the number involved in defining the upper bound for the negative eigenvalues, as the left end point of the positive interval $[a, b]$. This choice plays no important role in the convergence analysis that will be presented in section 6. One may define this left end point to be any fixed positive number $< b$. The crucial observation is that the number of eigenvalues that are larger than 2α and less than a is of order $O(\ln(\alpha^{-1}))$.

Krylov subspace solvers are known to handle problems with spectra of the form (4.18) very well. This will be illuminated by numerical experiments in the next section. The rigorous convergence analysis of the MINRES method, which is based on (4.18) and Chebyshev polynomials, is rather standard and very technical. We will return to this issue after the presentation of the examples.

5. Numerical experiments.

Example 1. Let $\Omega = (0,1) \times (0,1)$ denote the unit square with boundary $\partial\Omega$ and consider the following optimization task:

$$\min_{v \in L^2(\Omega),\, u \in H^1(\Omega)} \left\{ \frac{1}{2} \|Tu - d\|^2_{L^2(\partial\Omega)} + \frac{1}{2}\alpha \|v\|^2_{L^2(\Omega)} \right\} \tag{5.1}$$

subject to

$$-\Delta u + u = \begin{cases} -v & \text{in } D = (0.25, 0.75) \times (0.25, 0.75), \\ 0 & \text{in } \Omega \setminus D, \end{cases} \tag{5.2}$$

$$\nabla u \cdot \vec{n} = 0 \quad \text{on } \partial\Omega. \tag{5.3}$$

In this case

$$\begin{aligned} H_1 &= L^2(\Omega), \\ H_2 &= H^1(\Omega), \\ H_3 &= L^2(\partial\Omega). \end{aligned}$$

Furthermore, the observation operator T is simply the L^2-trace of $u \in H^1(\Omega)$:

$$T : H^1(\Omega) \to L^2(\partial\Omega), \quad u \to u|_{\partial\Omega}.$$

Algorithmic details. Note that the weak form of the state equation (5.2)–(5.3) reads

$$\widehat{A}u = -\widehat{B}v, \tag{5.4}$$

where

$$\widehat{A} : H_2 \to H_2', \quad u \to (u, \phi)_{H^1(\Omega)} \quad \text{for all } \phi \in H^1(\Omega), \tag{5.5}$$

$$\widehat{B} : H_1 \to H_2', \quad v \to (v, \phi)_{L^2(D)} \quad \text{for all } \phi \in H^1(\Omega), \tag{5.6}$$

and H_2' denotes the dual space of H_2. In the previous sections we assumed that the operator A on the left-hand side of the state equation (1.2) is a mapping from the state space onto the state space. Note that $\widehat{A}$ does not fulfill this criterion. This can be fixed by invoking the Riesz map R_2 of $H_2 = H^1(\Omega)$:

$$R_2 : H_2 \to H_2'.$$

More precisely, by applying R_2^{-1} to both sides of (5.4) we get

$$R_2^{-1}\widehat{A}u = -R_2^{-1}\widehat{B}v,$$

which is on the desired form (1.2) since

$$\begin{aligned} A &= R_2^{-1}\widehat{A} : H_2 \to H_2, \\ B &= R_2^{-1}\widehat{B} : H_1 \to H_2. \end{aligned}$$

Since $(R_2^{-1})^* = R_2$ (see Appendix B), it follows that

$$\begin{aligned} A^* &= (\widehat{A})^* (R_2^{-1})^* \\ &= R_2^{-1} R_2 (\widehat{A})^* R_2 \\ &= R_2^{-1} (\widehat{A})', \end{aligned}$$

where

$$(\widehat{A})' = R_2 (\widehat{A})^* R_2 : H_2 \to H_2'$$

is the dual operator of $\widehat{A}$; see Appendix C. Similarly, if

$$(\widehat{B})' = R_1 (\widehat{B})^* R_2 : H_2 \to H_1'$$

denotes the dual operator of $\widehat{B}$, and R_1 is the Riesz map of H_1, then

$$\begin{aligned} B^* &= (\widehat{B})^* (R_2^{-1})^* \\ &= R_1^{-1} R_1 (\widehat{B})^* R_2 \\ &= R_1^{-1} (\widehat{B})'. \end{aligned}$$

We can therefore write $\mathcal{A}_\alpha$, defined in (3.2), in the form

$$\begin{aligned} \mathcal{A}_\alpha &= \begin{bmatrix} \alpha I & 0 & B^* \\ 0 & T^*T & A^* \\ B & A & 0 \end{bmatrix} \\ &= \begin{bmatrix} \alpha I & 0 & R_1^{-1}(\widehat{B})' \\ 0 & T^*T & R_2^{-1}(\widehat{A})' \\ R_2^{-1}\widehat{B} & R_2^{-1}\widehat{A} & 0 \end{bmatrix} \\ &= \underbrace{\begin{bmatrix} R_1^{-1} & 0 & 0 \\ 0 & R_2^{-1} & 0 \\ 0 & 0 & R_2^{-1} \end{bmatrix}}_{=\mathcal{R}^{-1}} \underbrace{\begin{bmatrix} \alpha R_1 & 0 & (\widehat{B})' \\ 0 & R_2 T^*T & (\widehat{A})' \\ \widehat{B} & \widehat{A} & 0 \end{bmatrix}}_{\widehat{\mathcal{A}_\alpha}}. \end{aligned} \tag{5.7}$$

Please note that one may regard $\mathcal{R}^{-1}$ to be a preconditioner:

- The KKT system

$$\mathcal{A}_\alpha p = b$$

can be expressed as

$$\mathcal{R}^{-1}\widehat{\mathcal{A}_\alpha} p = \mathcal{R}^{-1}\widehat{b},$$

where

$$\widehat{b} = \mathcal{R}b = \begin{bmatrix} 0 \\ R_2 T^* d \\ 0 \end{bmatrix}; \tag{5.8}$$

cf. (3.2)–(3.3).

- For every $\alpha > 0$, $\widehat{\mathcal{A}_\alpha}$ is an isomorphism mapping $H_1 \times H_2 \times H_2$ onto $(H_1 \times H_2 \times H_2)'$. This follows from the Babuška–Brezzi conditions, which can be verified to hold by arguments very similar to those presented in Appendix A.
- $\mathcal{R}^{-1}$ is an isomorphism mapping the dual space $(H_1 \times H_2 \times H_2)'$ back to the Hilbert space $H_1 \times H_2 \times H_2$. Further details can be found in [24]. In fact, $\mathcal{R}^{-1}$ is the inverse of the Riesz map of $H_1 \times H_2 \times H_2$.

Software systems typically require that all linear operators be represented in terms of matrices. In the present example, $H_1 = L^2(\Omega)$ and $H_2 = H^1(\Omega)$, and the Riesz maps R_1 and R_2 thus yield the mass matrix $\bar{M}$ and the sum $\bar{L}$ of the mass matrix and the stiffness matrix, respectively. From (5.5)–(5.6) we find that also $\widehat{B}$ and $\widehat{A}$ will (cf. section 6 in [24]) lead to $\bar{M}$ and $\bar{L}$, respectively. The treatment of $R_2 T^* T$ is somewhat more involved, but the "end product" is the mass matrix $\bar{M}_{\partial\Omega}$ associated with the boundary of the domain $\Omega = (0,1) \times (0,1)$. The matrix product associated with (5.7) is therefore

$$\underbrace{\begin{bmatrix} \bar{M}^{-1} & 0 & 0 \\ 0 & \bar{L}^{-1} & 0 \\ 0 & 0 & \bar{L}^{-1} \end{bmatrix}}_{=\bar{\mathcal{R}}^{-1}} \underbrace{\begin{bmatrix} \alpha\bar{M} & 0 & \bar{M}^T \\ 0 & \bar{M}_{\partial\Omega} & \bar{L}^T \\ \bar{M} & \bar{L} & 0 \end{bmatrix}}_{=\bar{\mathcal{A}}_\alpha}. \tag{5.9}$$

Furthermore, $\widehat{b}$ in (5.8) yields the vector

$$\bar{b} = \begin{bmatrix} 0 \\ \bar{M}_{\partial\Omega} \bar{d} \\ 0 \end{bmatrix}, \tag{5.10}$$

where $\bar{d}$ denotes the vector with the observation data (on $\partial\Omega$). Hence, we obtain the algebraic system

$$\bar{\mathcal{R}}^{-1} \bar{\mathcal{A}}_\alpha \bar{p} = \bar{\mathcal{R}}^{-1} \bar{b}. \tag{5.11}$$

The operator norms of $\mathcal{R}^{-1}$ and $\widehat{\mathcal{A}_\alpha}$, defined in (5.7), are bounded independently of any mesh parameter $h > 0$. Consequently, if sound discretization techniques are employed, this desired property will be inherited by the associated matrix product (5.9), and one obtains schemes that are robust with respect to h; cf. [24].

In most practical situations it is inconvenient, due to computational demands, to explicitly use the matrices $\bar{M}^{-1}$ and $\bar{L}^{-1}$ as preconditioners. Instead one may employ approximations defined by, e.g., multigrid cycles. Please note the following:

- The iteration counts presented below were produced with approximations of $\bar{M}^{-1}$ and $\bar{L}^{-1}$ consisting of one call of algebraic multigrid (AMG) with symmetric successive overrelaxation (SSOR) as smoother; see [23] for a description of the software framework cbc.block, built on top of FEniCS and PyTrilinos. If not stated otherwise, the AMG call consisted of one V-cycle with two SSOR sweeps. One AMG call was made per MINRES iteration.

- All figures in this section showing eigenvalue distributions of $\bar{\mathcal{R}}^{-1}\bar{\mathcal{A}}_\alpha$ were generated with the true inverse matrices $\bar{M}^{-1}$ and $\bar{L}^{-1}$.

The optimality system was discretized with the standard finite element method using piecewise linear basis functions. More specifically, the domain $\Omega = (0,1)\times(0,1)$ was divided in $N \times N$ squares, and each of these squares was split into two triangles.

All iteration counts presented in this paper were generated with a true solution $\bar{p}^* = 0$ and a random initial guess $\bar{p}_0$. The iteration process was stopped as soon as

$$\frac{\|\bar{r_k}\|}{\|\bar{r_0}\|} = \left[\frac{(\bar{\mathcal{A}}_\alpha\bar{p}_k - \bar{b}, \tilde{\mathcal{R}}^{-1}\{\bar{\mathcal{A}}_\alpha\bar{p}_k - \bar{b}\})}{(\bar{\mathcal{A}}_\alpha\bar{p}_0 - \bar{b}, \tilde{\mathcal{R}}^{-1}\{\bar{\mathcal{A}}_\alpha\bar{p}_0 - \bar{b}\})}\right]^{1/2} < \epsilon, \tag{5.12}$$

where $\epsilon = 10^{-6}$ or $\epsilon = 10^{-10}$. Furthermore, $\bar{r_k} = \bar{\mathcal{A}}_\alpha p_k - \bar{b}$ is the residual vector, and $\bar{p}_k$ is the kth approximation of $\bar{p}^*$ generated by the MINRES algorithm. In the stopping criterion (5.12), $\tilde{\mathcal{R}}^{-1}$ is the multigrid approximation of the matrix $\bar{\mathcal{R}}^{-1}$ defined in (5.9).

Results. Tables 1 and 2 show that the number of MINRES iterations needed to solve (5.11) does not increase dramatically as α decreases or as N increases. For example, the numbers in the last row in Table 1 are modeled well by the formula

$$42.5 - 6.5\log_{10}(\alpha).$$

(We used the method of least squares to compute the constants in this expression.) Similarly, the iteration counts presented in the last row in Table 2 are rather well represented by the expression

$$71 - 27\log_{10}(\alpha).$$

We also explored how the number of SSOR sweeps in the AMG preconditioner influenced the convergence behavior; see Table 3. Performing two SSOR sweeps clearly reduces the required number of iterations compared with only applying one sweep.

TABLE 1

Number of MINRES iterations required to solve the model problem studied in Example 1 *with* $\epsilon = 10^{-6}$ *in* (5.12). *(For* $\alpha \le 0.001$ *the stopping criterion* $\epsilon = 10^{-6}$ *was not strict enough, i.e., the approximate solution contained a significant amount of error.)*

$N\backslash\alpha$	1.0	0.1	0.01
32	40	49	50
64	33	38	44
128	33	38	44
256	47	52	58
512	43	48	56

TABLE 2

Number of MINRES iterations required to solve the model problem studied in Example 1 *with* $\epsilon = 10^{-10}$ *in* (5.12).

$N\backslash\alpha$	1.0	0.1	0.01	0.001	0.0001
32	54	65	79	115	132
64	52	65	84	116	140
128	55	69	93	112	140
256	71	89	117	135	180
512	77	92	126	147	186

TABLE 3

The table shows how the number of SSOR sweeps in the AMG preconditioner influences the convergence behavior for the model problem studied in Example 1. *These numbers were produced with* $\alpha = 0.0001$ *and* $\epsilon = 10^{-10}$ *in* (5.12).

N	32	64	128	256	512
1 sweep	206	185	191	263	277
2 sweeps	131	141	139	180	159

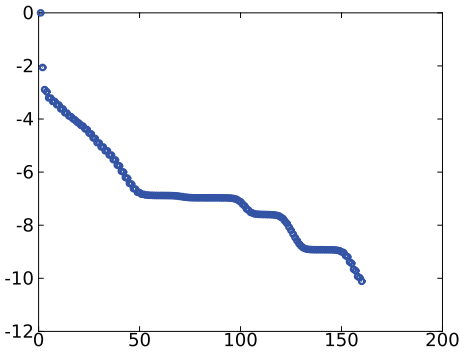

FIG. 1. *A typical convergence curve associated with the experiments presented in Example* 1, *that is,* $\log_{10}(\|\bar{r_k}\|/\|\bar{r_0}\|)$ *as a function of the iteration number* k. *Here,* $N = 512$ *and* $\alpha = 0.0001$.

On the other hand, the convergence criterion (5.12) depends on the number of SSOR sweeps because the matrix $\tilde{\mathcal{R}}^{-1}$ does. A more thorough investigation of this issue is therefore needed. Especially, one should seek to estimate the optimal number of sweeps, but this is beyond the scope of the present text.

Figure 1 contains the convergence curve, i.e., $\log_{10}(\|\bar{r_k}\|/\|\bar{r_0}\|)$ as a function of the iteration number k, for the case $N = 512$ and $\alpha = 0.0001$. We see that the curve goes rapidly down to $\approx 10^{-7}$, then "stalls" for quite a large number of iterations, and finally moves nicely down to 10^{-10}. We do not have any good explanation for the lack of linear descent of $\log_{10}(\|\bar{r_k}\|/\|\bar{r_0}\|)$ in some intervals.

The left panel of Figure 2 shows the eigenvalues[3] of $\mathcal{A}_\alpha$ sorted in increasing order. The three intervals used to characterize the spectrum of $\mathcal{A}_\alpha$ in (4.18) in Theorem 4.5 are clearly visible. Note that the spectrum contains only a handful of isolated eigenvalues; see the right panel of Figure 2.

In our analysis we assumed that $\mathcal{A}_0$ inherits the ill posed nature of the optimization problem (1.1)–(1.2); see assumption $\mathcal{A}6$. According to Figure 3 this is indeed the case for the present model problem. Moreover, the right panel of Figure 3 indicates that $\mathcal{A}_0$ satisfies assumption $\mathcal{A}6$, i.e., the problem is severely ill posed.

Example 2. Example 2 is identical to Example 1, but we introduce a variable coefficient in the state equation:

$$-\nabla\cdot(\gamma\nabla u) + u = \begin{cases} -v & \text{in } D = (0.25, 0.75)\times(0.25, 0.75), \\ 0 & \text{in } \Omega\setminus D, \end{cases}$$
$$\gamma(x, y) = 2 + \sin(2\pi(x + y)), \quad (x, y) \in (0, 1).$$

[3]Recall that $\mathcal{A}_\alpha = \mathcal{R}^{-1}\widehat{\mathcal{A}_\alpha}$ (see (5.7)), with associated matrix representation $\bar{\mathcal{R}}^{-1}\bar{\mathcal{A}}_\alpha$; cf. (5.9). One can show that $\mathcal{A}_\alpha$ and $\bar{\mathcal{R}}^{-1}\bar{\mathcal{A}}_\alpha$ are similar operators and therefore have the same eigenvalues.

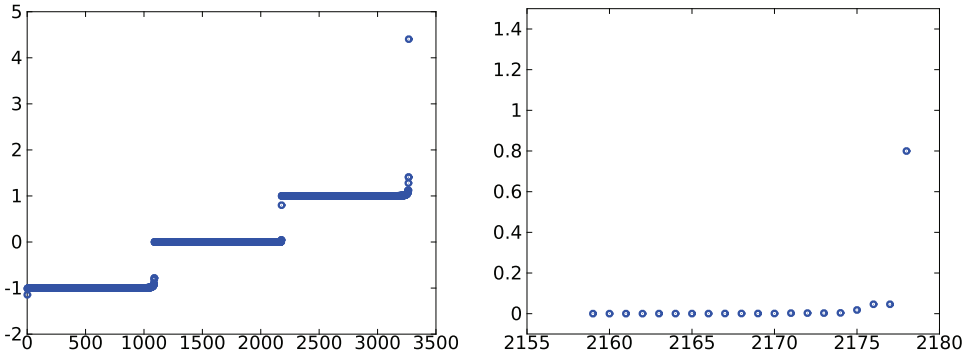

FIG. 2. *The left panel shows the eigenvalues of $\mathcal{A}_\alpha$, sorted in increasing order, for the model problem studied in Example* 1. *These numbers were generated with $\alpha = 0.0001$ and $N = 32$. For this problem the constants in Lemma* 4.1 *are $c \approx 0.981$ and $C \approx 4.405$. In the right panel we have zoomed in on the isolated eigenvalues that are larger than 2α and less than $a = 1$; see Theorem* 4.5.

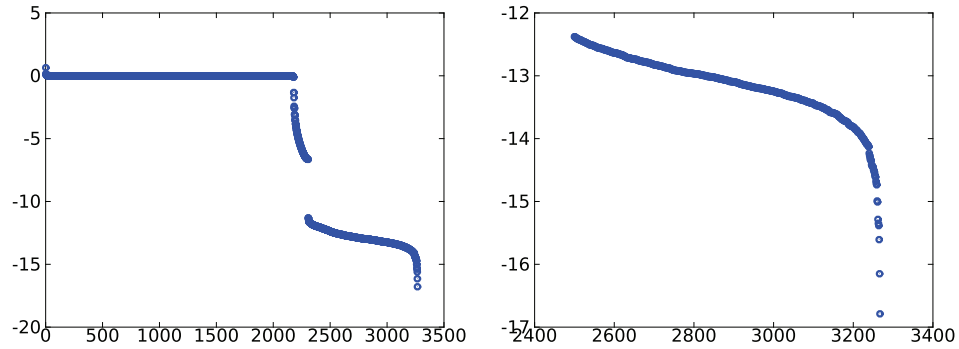

FIG. 3. *The left panel shows the logarithm of the absolute value of the eigenvalues of $\mathcal{A}_0$ sorted in decreasing order. These are results obtained in Example* 1 *with $N = 32$. In the right panel we have zoomed in on the interval* $(2500, 3267)$.

This change implies that the matrix $\bar{L}$ in $\bar{\mathcal{A}}_\alpha$ in (5.9) must be replaced with the matrix associated with the operator $-\nabla \cdot (\gamma \nabla u) + u$. We would like to emphasize that the same preconditioner is employed in Examples 1 and 2.

Tables 4 and 5 contain the iteration counts for this model problem. Concerning the influence of α and N, we see that the conclusions reached in Example 1 are still valid, but more iterations are required due to the variable coefficient. Please observe that the iteration counts in the last row in Table 4 are approximately of the form

$$93 - 18 \log_{10}(\alpha).$$

A similar property holds for the numbers in the last row in Table 5. That is, these numbers are well represented by the formula

$$163 - 45 \log_{10}(\alpha).$$

Information about the eigenvalue distributions of $\mathcal{A}_\alpha$ and $\mathcal{A}_0$ can be found in Figures 4 and 5, respectively. The change in the spectrum caused by the variable coefficient is clearly visible in Figure 4, which should be compared with Figure 2. This observation is consistent with the numbers presented in Tables 4 and 1. Also note that $\mathcal{A}_\alpha$ has a few isolated eigenvalues, see the right panel of Figure 4.

TABLE 4

Number of MINRES iterations required to solve the model problem studied in Example 2 *with* $\epsilon = 10^{-6}$ *in* (5.12). *(For* $\alpha \leq 0.001$ *the stopping criterion* $\epsilon = 10^{-6}$ *was not strict enough; i.e., the approximate solution contained a significant amount of error.)*

$N\backslash\alpha$	1.0	0.1	0.01
32	84	98	106
64	78	92	101
128	75	90	104
256	103	119	135
512	95	108	131

TABLE 5

Number of MINRES iterations required to solve the model problem studied in Example 2 *with* $\epsilon = 10^{-10}$ *in* (5.12).

$N\backslash\alpha$	1.0	0.1	0.01	0.001	0.0001
32	116	140	172	249	302
64	115	141	180	254	316
128	124	151	200	278	304
256	161	193	247	283	392
512	165	198	259	305	336

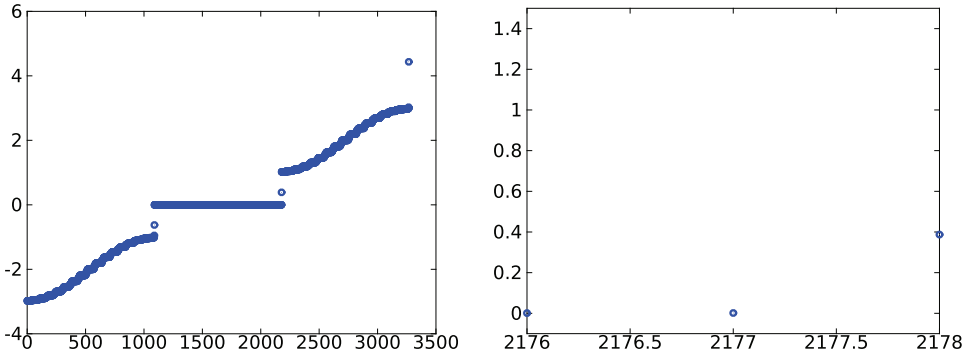

FIG. 4. *The left panel shows the eigenvalues of* $\mathcal{A}_\alpha$, *sorted in increasing order, for the model problem studied in Example* 2. *These numbers were generated with* $\alpha = 0.0001$ *and* $N = 32$. *For this problem the constants in Lemma* 4.1 *are* $c \approx 0.998$ *and* $C \approx 4.436$. *In the right panel we have zoomed in on the isolated eigenvalues that are larger than* 2α *and less than* $a = 1$; *see Theorem* 4.5.

By inspecting Figures 5 and 3 we conclude that the qualitative properties of the small eigenvalues, in the absolute sense, of $\mathcal{A}_0$ are not significantly influenced by the coefficient function γ. In fact, the right panel in Figure 5 indicates that the present problem is severely ill posed.

Example 3. Our third test case is the "standard" test problem of the PDE constrained optimization community:

$$\min_{v \in L^2(\Omega),\, u \in H^1(\Omega)} \left\{ \frac{1}{2} \|Tu - d\|^2_{L^2(\Omega)} + \frac{1}{2} \alpha \|v\|^2_{L^2(\Omega)} \right\} \tag{5.13}$$

subject to

$$-\Delta u + u = -v \quad \text{in } \Omega, \tag{5.14}$$

$$\nabla u \cdot \vec{n} = 0 \quad \text{on } \partial\Omega. \tag{5.15}$$

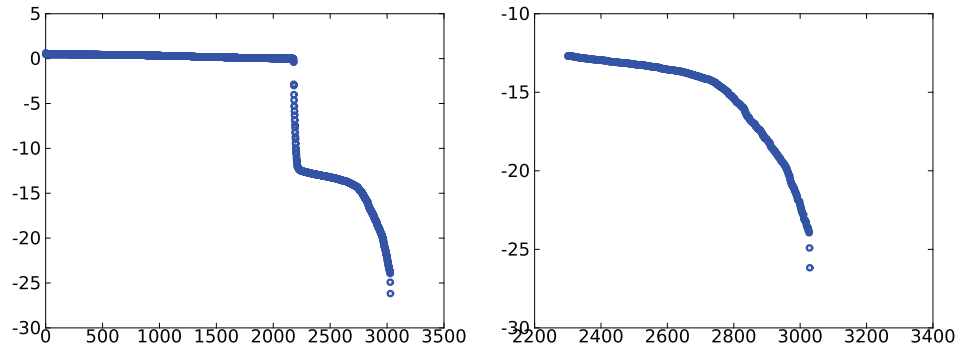

FIG. 5. *The left panel shows the logarithm of the absolute value of the eigenvalues of $\mathcal{A}_0$ sorted in decreasing order. These are results obtained in Example 2 with $N = 32$. In the right panel we have zoomed in on the interval $(2300, 3267)$; some of the eigenvalues are not displayed because they are zero (the logarithm of zero is not defined).*

TABLE 6
Number of MINRES iterations required to solve the model problem studied in Example 3 with $\epsilon = 10^{-10}$ in (5.12).

$N\backslash\alpha$	1.0	0.1	0.01	0.001	0.0001
32	51	66	78	121	188
64	48	64	86	127	203
128	52	68	88	137	212
256	68	89	108	163	232
512	72	92	118	173	215

We observe that

$$H_1 = L^2(\Omega),$$
$$H_2 = H^1(\Omega),$$
$$H_3 = L^2(\Omega)$$

and that the observation operator T is the imbedding

$$T : H^1(\Omega) \hookrightarrow L^2(\Omega), \quad u \to u.$$

The latter fact implies that $\bar{M}_{\partial\Omega}$ in $\bar{\mathcal{A}}_\alpha$ and $\bar{b}$ (see (5.9) and (5.10)) must be replaced with the mass matrix $\bar{M}$.

Table 6 shows that the MINRES method also solves this problem efficiently. More specifically, there is no dramatic increase in the number of iterations needed as α decreases or N increases. We observe that the iteration counts in the last row are quite well modeled by the formula

$$61 - 37\log_{10}(\alpha).$$

The eigenvalues associated with our third model problem is depicted in Figure 6. Again, we observe that the spectrum mainly consists of three bounded intervals and a small number of isolated eigenvalues.

Figure 7 contains log-log plots of the absolute value of the eigenvalues of $\mathcal{A}_0$ sorted in decreasing order. Note that the right panel almost shows a line, and that this is log-log plot. This indicates that (5.13)–(5.15) is mildly ill posed: For a mildly ill

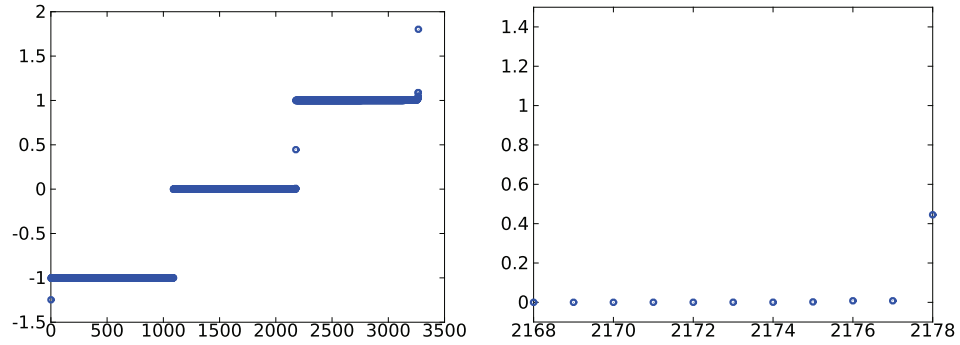

FIG. 6. *The left panel shows the eigenvalues of $\mathcal{A}_\alpha$, sorted in increasing order, for the model problem studied in Example* 3. *These numbers were generated with $\alpha = 0.0001$ and $N = 32$. For this problem the constants in Lemma* 4.1 *are $c \approx 1.0$ and $C \approx 1.802$. In the right panel we have zoomed in on the isolated eigenvalues that are larger than 2α and less than $a = 1$; see Theorem* 4.5.

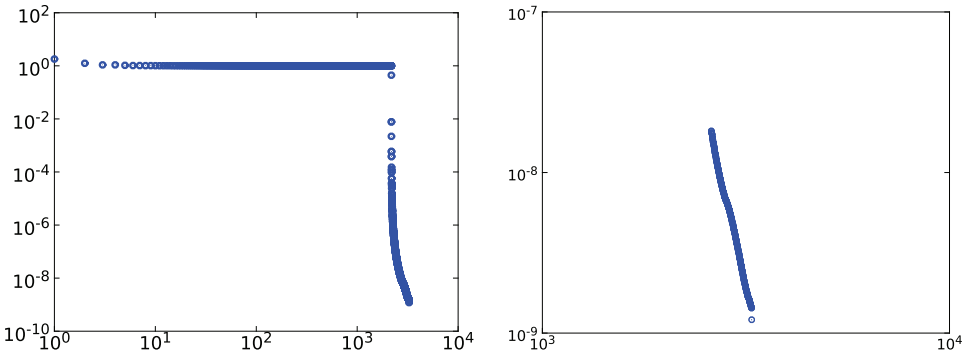

FIG. 7. *The left panel shows a log-log plot of the absolute value of the eigenvalues of $\mathcal{A}_0$ sorted in decreasing order. These are results obtained in Example* 3 *with $N = 32$. In the right panel we have zoomed in on the interval $(10^{3.41}, 10^{3.51})$.*

posed problem the eigenvalues of $\mathcal{A}_0$ would be expected to obey

$$|\lambda_i(\mathcal{A}_0)| \leq C\, i^{-c} \quad \text{for } i = 1, 2, \ldots, \tag{5.16}$$

and hence

$$\ln(|\lambda_i(\mathcal{A}_0)|) \leq \ln(C) - c \ln(i).$$

In the theoretical considerations presented in section 4 we assumed that the inverse problem (1.1)–(1.2) is severely ill posed and that this property is inherited by $\mathcal{A}_0$ in the sense of assumption $\mathcal{A}6$. If instead (1.1)–(1.2) is mildly ill posed, then one would expect (5.16) to hold. The analysis presented in the previous sections can easily be modified to also include such cases. The only significant change concerns the number $N(\alpha)$ of isolated eigenvalues in Theorem 4.5:

$$N(\alpha) \leq \left\lceil \left(C\alpha^{-1}\right)^{1/c} \right\rceil = O(\alpha^{-1/c}).$$

We conclude, at least from an asymptotic point of view, that severely ill posed problems will have fewer isolated eigenvalues than mildly ill posed problems. Consequently,

one would expect that the required number of MINRES iterations will increase faster for mildly ill posed problems than for severely ill posed problems as $\alpha \to 0$. This is partially observed in our examples:

- In the last row of Table 6 (Example 3, mildly ill posed) the number of iterations grows from 72 to 215, i.e., by a factor of $215/72 \approx 2.99$.
- The corresponding factors associated with Tables 2 (Example 1, severely ill posed) and 5 (Example 2, severely ill posed) are $186/77 \approx 2.42$ and $336/165 \approx 2.04$, respectively.

6. Convergence analysis. We know that the spectrum of $\mathcal{A}_\alpha$ consists of three bounded intervals and a few isolated eigenvalues. As will become evident below, the convergence analysis based on this observation and Chebyshev polynomials gets very involved. Such arguments are much simpler for systems with spectra contained in one bounded interval and with a finite number of isolated eigenvalues outside this interval; see Axelsson and Lindskog [5] and Axelsson [3].

In this section we proceed as follows:

- First, we present the theorem.
- Thereafter we discuss the numerical experiments, presented above, in view of this result.
- Finally, the proof of the theorem is presented.

Our main concern is the number of MINRES iterations required to solve

$$\mathcal{A}_\alpha p = b. \tag{6.1}$$

Before we formulate our last result, please recall the basic structure of the spectrum of $\mathcal{A}_\alpha$:

$$\operatorname{sp}(\mathcal{A}_\alpha) \subset [-b,-a] \cup [c\alpha, 2\alpha] \cup \{\lambda_1, \lambda_2, \dots, \lambda_{N(\alpha)}\} \cup [a,b]; \tag{6.2}$$

see Theorem 4.5.

THEOREM 6.1. *Let p^* denote the solution of* (6.1) *and let $\epsilon > 0$ be a given error tolerance. If*

$$\begin{aligned} &k \geq \frac{b}{a}\left[n \ln\left(\frac{4b^2}{(4-c^2)\alpha^2}\right) + N(\alpha)\ln\left(\frac{b^2}{4\alpha^2}\right) + \ln(2)\right] + 2n + 2N(\alpha) + 2, \\ &\text{where } n = \left\lceil \frac{1}{c}\ln\left(\frac{2}{\epsilon}\right)\right\rceil, \end{aligned} \tag{6.3}$$

then

$$\frac{\|p_k - p^*\|}{\|p_0 - p^*\|} = \left(\frac{(\mathcal{A}_\alpha(p_k - p^*), \mathcal{A}_\alpha(p_k - p^*))}{(\mathcal{A}_\alpha(p_0 - p^*), \mathcal{A}_\alpha(p_0 - p^*))}\right)^{1/2} \leq \epsilon, \tag{6.4}$$

provided that

$$\begin{aligned} &\alpha \in (0, \alpha^*], \\ &\alpha^* = \min\{1, a/2\}. \end{aligned}$$

Here, p_k is the kth approximation of p^* generated by the MINRES method applied to (6.1)*, and $N(\alpha)$ is of order $O(\ln(\alpha^{-1}))$:*

$$N(\alpha) \leq \left\lceil \frac{\ln(\tilde{c}) - \ln(\alpha)}{\tilde{C}} \right\rceil = O(\ln(\alpha^{-1})).$$

Remark. If the state equation is a PDE, then recall that $\mathcal{A}_\alpha = \mathcal{R}^{-1}\widehat{\mathcal{A}_\alpha}$ (see (5.7)), where $\mathcal{R}^{-1}$ is the inverse of the Riesz map of $H_1 \times H_2 \times H_2$. Hence, for all $z \in H_1 \times H_2 \times H_2$,

$$\begin{aligned}(\mathcal{A}_\alpha z, \mathcal{A}_\alpha z) &= (\mathcal{R}^{-1}\widehat{\mathcal{A}_\alpha} z, \mathcal{R}^{-1}\widehat{\mathcal{A}_\alpha} z) \\ &= \langle \widehat{\mathcal{A}_\alpha} z, \mathcal{R}^{-1}\widehat{\mathcal{A}_\alpha} z \rangle.\end{aligned}$$

This verifies that (6.4) corresponds to the standard stopping criterion used in connection with the preconditioned MINRES algorithm.

Since the number $N(\alpha)$ of isolated eigenvalues is of order $O(\ln(\alpha^{-1}))$, (6.3) shows that the required number of MINRES iterations cannot grow faster than $O([\ln(\alpha^{-1})]^2)$ as $\alpha \to 0$. However, for the numerical experiments discussed in section 5 we observed iteration counts that are close to order $O(\ln(\alpha^{-1}))$. Hence, at first glance, there seems to be some sort of discrepancy between Theorem 6.1 and our practical experience.

Note that the bound in (6.3) can be written in the form

$$q_0 + q_1 \ln(\alpha^{-1}) + 2N(\alpha) \ln(\alpha^{-1}).$$

We did not find many isolated eigenvalues in the examples studied above; see Figures 2, 4, and 6. It is therefore to be expected that $2N(\alpha)$ is significantly smaller than q_1, which would explain our observations.

Remark. There is an even more subtle aspect of this issue. If one reads the analysis in section 4 carefully, one finds that the constant 2 used to define the right end point of the interval $[c\alpha, 2\alpha]$ in (6.2) is arbitrary. In fact, 2 can be replaced by any positive number r larger than c, leading to an interval of the form $[c\alpha, r\alpha]$. Furthermore, choosing r larger will reduce the number of isolated eigenvalues—with the price that some of the other constants in (6.3) increase. One therefore gets an upper bound for the required number of iterations in the form

$$\min_r \left\{ q_0(r) + q_1(r) \ln(\alpha^{-1}) + 2N(\alpha, r) \ln(\alpha^{-1}) \right\},$$

where $N(\alpha, r)$ decreases as r grows.

Let us consider Theorem 6.1 for problems with a state equation (1.2) given in terms of a discretized PDE. In such cases, will the quantities in (6.3) depend on the grid parameter h? The constants a and b are defined in (4.16) and (4.17), respectively, where $[-\tilde{b}, -\tilde{a}]$ is the interval that contains the negative eigenvalues of $\mathcal{A}_\alpha$, and C is the upper bound (4.1) of the operator norm of $\mathcal{A}_\alpha$. Assume that the discretization of the PDE is sound, and that the approximate inverse Riesz map is applied as preconditioner. Then one can verify that the interval $[-\tilde{b}, -\tilde{a}]$ and C do not depend on the mesh parameter h. In a similar fashion one can establish that the constant c, used to define the interval $[c\alpha, 2\alpha]$ containing the small eigenvalues of $\mathcal{A}_\alpha$, is independent of h. In fact, c is merely the constant used in Lemma 4.1.

The question of whether the number $N(\alpha)$ of isolated eigenvalues depends on the mesh parameter h is more subtle. Assume that the PDE constrained optimization problem at hand is severely ill posed. Then, before discretization, inequality (3.5) in assumption $\mathcal{A}6$ will hold with constants c and C not depending on h. Upon discretization, it is therefore to be expected that this property is inherited by the discretized KKT system, i.e., that (3.5) is satisfied with constants that are independent of the mesh parameter. Consequently, according to the analysis presented in section 4.3, we can conclude that $N(\alpha)$ does not increase as h decreases. Please note that this is in accordance with our numerical experiments, presented in section 5, since the number of iterations does not change dramatically on fine meshes.

Proof of Theorem 6.1. Theorem 4.5 states that

$$\begin{aligned}&\mathrm{sp}(\mathcal{A}_\alpha) \subset [-b,-a] \cup [c\alpha, 2\alpha] \cup \{\lambda_1, \lambda_2, \ldots, \lambda_{N(\alpha)}\} \cup [a,b],\\ &2\alpha < \lambda_i < a \quad \text{for } i = 1,2,\ldots,N(\alpha).\end{aligned}$$

The squares of the eigenvalues must therefore satisfy

$$\begin{aligned}&\lambda^2 \in [c^2\alpha^2, 4\alpha^2] \cup \{\lambda_1^2, \lambda_2^2, \ldots, \lambda_{N(\alpha)}^2\} \cup [a^2, b^2],\\ &4\alpha^2 < \lambda_i^2 < a^2 \quad \text{for } i = 1,2,\ldots,N(\alpha).\end{aligned} \tag{6.5}$$

An analysis of the conjugate gradient method applied to positive definite systems with spectra contained in two intervals is presented on pages 19–21 in Axelsson [3]. It is possible to adapt Axelsson's argument to the present situation. The main challenge is to incorporate the effect of the isolated eigenvalues.

According to Elman, Silvester, and Wathen [12, page 306],

$$\frac{\|p_k - p^*\|}{\|p_0 - p^*\|} \leq \min_{\Phi_k \in \overline{\Pi}_k} \max_{\lambda \in \mathrm{sp}(\mathcal{A}_\alpha)} |\Phi_k(\lambda)|, \tag{6.6}$$

where $\overline{\Pi}_k$ is the set of all polynomials of degree $\leq k$ with $\Phi_k(0) = 1$.

We now aim at constructing a suitable polynomial $\Psi_k \in \overline{\Pi}_k$. To this end, let T_s denote the Chebyshev polynomial of order s. Consider

$$\begin{aligned}&\Phi_m^*(x;\, c^2\alpha^2, 4\alpha^2) = \frac{T_m\left(\frac{4\alpha^2 + c^2\alpha^2 - 2x}{4\alpha^2 - c^2\alpha^2}\right)}{T_m\left(\frac{4\alpha^2 + c^2\alpha^2}{4\alpha^2 - c^2\alpha^2}\right)},\\ &P_{N(\alpha)}(x) = \prod_{i=1}^{N(\alpha)} \left(1 - \frac{x}{\lambda_i^2}\right),\\ &\Phi_{l-q/2}^*(x;\, a^2, b^2) = \frac{T_{l-q/2}\left(\frac{b^2 + a^2 - 2x}{b^2 - a^2}\right)}{T_{l-q/2}\left(\frac{b^2 + a^2}{b^2 - a^2}\right)},\\ &q = 2m + 2N(\alpha),\\ &l = \lceil k/2 \rceil - 1,\end{aligned}$$

where k is the number of MINRES iterations and m is a positive integer that will be specified below. We suggest employing the following polynomial in the convergence analysis:

$$\Psi_k(\lambda) = \Phi_m^*(\lambda^2;\, c^2\alpha^2, 4\alpha^2) \cdot P_{N(\alpha)}(\lambda^2) \cdot \Phi_{l-q/2}^*(\lambda^2;\, a^2, b^2), \tag{6.7}$$

which has degree

$$2m + 2N(\alpha) + 2l - q \leq 2m + 2N(\alpha) + k - 2m - 2N(\alpha) = k$$

and $\Psi_k(0) = 1$, i.e., $\Psi_k \in \overline{\Pi}_k$.

Our goal is to determine a suitable upper bound for

$$|\Psi_k(\lambda)| \quad \text{for } \lambda \in [-b,-a] \cup [c\alpha, 2\alpha] \cup \{\lambda_1, \lambda_2, \ldots, \lambda_{N(\alpha)}\} \cup [a,b];$$

see (6.6). This can be accomplished as follows:

- *Assume that* $\lambda \in \{\lambda_{\mathbf{1}}, \lambda_{\mathbf{2}}, \ldots, \lambda_{\mathbf{N}(\alpha)}\}$. Then $P_{N(\alpha)}(\lambda^2) = 0$, which implies that

$$\underline{\Psi_k(\lambda) = 0.}$$

- *Assume that* $\lambda \in [\mathbf{c}\alpha, \mathbf{2}\alpha] \Rightarrow \lambda^{\mathbf{2}} \in [\mathbf{c^2}\alpha^{\mathbf{2}}, \mathbf{4}\alpha^{\mathbf{2}}]$. We treat each of the three factors in (6.7) separately:
  - Because $\lambda^2 \leq 4\alpha^2 < \lambda_i^2$ (see (6.5)), we find that

$$\left|1 - \frac{\lambda^2}{\lambda_i^2}\right| < 1 \Rightarrow |P_{N(\alpha)}(\lambda^2)| < 1.$$

  - Since $\lambda^2 < a^2 < b^2$, it follows that

$$1 < \frac{b^2 + a^2 - 2\lambda^2}{b^2 - a^2} < \frac{b^2 + a^2}{b^2 - a^2},$$

    and by well-known properties of Chebyshev polynomials

$$\left|T_{l-q/2}\left(\frac{b^2 + a^2 - 2\lambda^2}{b^2 - a^2}\right)\right| < \left|T_{l-q/2}\left(\frac{b^2 + a^2}{b^2 - a^2}\right)\right|$$
$$\Downarrow$$
$$|\Phi^*_{l-q/2}(\lambda^2;\, a^2, b^2)| < 1.$$

  - It is well known (see, e.g., Axelsson and Lindskog [5] and the references therein) that

$$\max_{\lambda \in [c\alpha, 2\alpha]} |\Phi^*_m(\lambda^2;\, c^2\alpha^2, 4\alpha^2)| \leq 2\left(\frac{1 - \sqrt{\frac{c^2\alpha^2}{4\alpha^2}}}{1 + \sqrt{\frac{c^2\alpha^2}{4\alpha^2}}}\right)^m$$
$$= 2\left(\frac{1 - \frac{c}{2}}{1 + \frac{c}{2}}\right)^m. \tag{6.8}$$

It therefore is evident that, for $\lambda \in [c\alpha, 2\alpha]$,

$$|\Psi_k(\lambda)| \leq 2\left(\frac{1 - \frac{c}{2}}{1 + \frac{c}{2}}\right)^m.$$

We can now specify the (thus far) unspecified positive integer m from the criterion

$$|\Psi_k(\lambda)| \leq \epsilon,$$

which yields

$$m = \left\lceil \frac{1}{c} \ln\left(\frac{2}{\epsilon}\right) \right\rceil. \tag{6.9}$$

- *Assume that* $\lambda \in [-b,-a]\cup[a,b] \Rightarrow \lambda^2 \in [a^2,b^2]$. The three factors of Ψ_k (see (6.7)) are first analyzed individually:
  - Note that
$$\left|1-\frac{\lambda^2}{\lambda_i^2}\right| = \left|\frac{\lambda_i^2-\lambda^2}{\lambda_i^2}\right| \le \frac{b^2}{4\alpha^2},$$
because $4\alpha^2 < \lambda_i^2 < \lambda^2 \le b^2$. Consequently,
$$|P_{N(\alpha)}(\lambda^2)| \le \left(\frac{b^2}{4\alpha^2}\right)^{N(\alpha)}. \tag{6.10}$$
  - Similarly to (6.8),
$$\begin{aligned}|\Phi^*_{l-q/2}(\lambda^2;\, a^2,b^2)| &\le 2\left(\frac{1-\sqrt{\frac{a^2}{b^2}}}{1+\sqrt{\frac{a^2}{b^2}}}\right)^{l-q/2}\\ &= 2\left(\frac{1-\frac{a}{b}}{1+\frac{a}{b}}\right)^{l-q/2}. \end{aligned}\tag{6.11}$$
  - The treatment of
$$\Phi^*_m(\lambda^2;\, c^2\alpha^2, 4\alpha^2) = \frac{T_m\left(\frac{4\alpha^2+c^2\alpha^2-2\lambda^2}{4\alpha^2-c^2\alpha^2}\right)}{T_m\left(\frac{4\alpha^2+c^2\alpha^2}{4\alpha^2-c^2\alpha^2}\right)}$$
is more involved. Since $\lambda^2 > 4\alpha^2 > c\alpha^2$, we find that
$$\left|\frac{4\alpha^2+c^2\alpha^2-2\lambda^2}{4\alpha^2-c^2\alpha^2}\right| > 1.$$
Chebyshev polynomials are known to satisfy
$$|T_m(y)| \le |2y|^m \quad \text{for } |y|>1$$
(see, e.g., page 20 in [3]), and we therefore conclude that
$$\begin{aligned}\left|T_m\left(\frac{4\alpha^2+c^2\alpha^2-2\lambda^2}{4\alpha^2-c^2\alpha^2}\right)\right| &\le \left|2\cdot\frac{4\alpha^2+c^2\alpha^2-2\lambda^2}{4\alpha^2-c^2\alpha^2}\right|^m\\ &\le \left|2\cdot\frac{4\alpha^2+c^2\alpha^2-2b^2}{(4-c^2)\alpha^2}\right|^m\\ &\le \left(\frac{4b^2}{(4-c^2)\alpha^2}\right)^m.\end{aligned}$$
If this bound is combined with the inequality
$$\left|\frac{1}{T_m\left(\frac{4\alpha^2+c^2\alpha^2}{4\alpha^2-c^2\alpha^2}\right)}\right| \le 2\left(\frac{1-\frac{c}{2}}{1+\frac{c}{2}}\right)^m$$
(see page 13 in [3]), then we can conclude that
$$|\Phi^*_m(\lambda^2;\, c^2\alpha^2, 4\alpha^2)| \le 2\left(\frac{1-\frac{c}{2}}{1+\frac{c}{2}}\right)^m\left(\frac{4b^2}{(4-c^2)\alpha^2}\right)^m. \tag{6.12}$$

For $\lambda \in [-b,-a] \cup [a,b]$ we thus conclude from (6.10), (6.11), and (6.12) that

$$|\Psi_k(\lambda)| \le 4 \left(\frac{1-\frac{c}{2}}{1+\frac{c}{2}}\right)^m \left(\frac{4b^2}{(4-c^2)\alpha^2}\right)^m \left(\frac{1-\frac{a}{b}}{1+\frac{a}{b}}\right)^{l-q/2} \left(\frac{b^2}{4\alpha^2}\right)^{N(\alpha)},$$

where

$$\begin{aligned} l &= \lceil k/2 \rceil - 1, \\ q &= 2m + 2N(\alpha), \\ m &= \left\lceil \frac{1}{c} \ln\left(\frac{2}{\epsilon}\right) \right\rceil. \end{aligned} \tag{6.13}$$

Let us summarize our findings: The polynomial Ψ_k, defined in (6.7), satisfies

- $\Psi_k(\lambda) = 0 \quad \text{for } \lambda \in \{\lambda_1, \lambda_2, \dots, \lambda_{N(\alpha)}\}$,
- $|\Psi_k(\lambda)| \le 2\big(\frac{1-\frac{c}{2}}{1+\frac{c}{2}}\big)^m$ for $\lambda \in [c\alpha, 2\alpha]$,
- $|\Psi_k(\lambda)| \le 4\big(\frac{1-\frac{c}{2}}{1+\frac{c}{2}}\big)^m \big(\frac{4b^2}{(4-c^2)\alpha^2}\big)^m \big(\frac{1-\frac{a}{b}}{1+\frac{a}{b}}\big)^{l-q/2} \big(\frac{b^2}{4\alpha^2}\big)^{N(\alpha)}$ for $\lambda \in [-b,-a] \cup [a,b]$.

If m is defined as in (6.13), then

$$2\left(\frac{1-\frac{c}{2}}{1+\frac{c}{2}}\right)^m \le \epsilon$$

and

$$|\Psi_k(\lambda)| \le 2\epsilon \left(\frac{4b^2}{(4-c^2)\alpha^2}\right)^m \left(\frac{1-\frac{a}{b}}{1+\frac{a}{b}}\right)^{l-q/2} \left(\frac{b^2}{4\alpha^2}\right)^{N(\alpha)}$$

for $\lambda \in [-b,-a] \cup [a,b]$, where

$$\begin{aligned} l &= \lceil k/2 \rceil - 1, \\ q &= 2m + 2N(\alpha). \end{aligned}$$

Consequently, for all $\lambda \in [-b,-a] \cup [c\alpha, 2\alpha] \cup \{\lambda_1, \lambda_2, \dots, \lambda_{N(\alpha)}\} \cup [a,b]$,

$$|\Psi_k(\lambda)| \le \epsilon,$$

provided that

$$\begin{aligned} k &\ge \frac{b}{a}\left[m \ln\left(\frac{4b^2}{(4-c^2)\alpha^2}\right) + N(\alpha)\ln\left(\frac{b^2}{4\alpha^2}\right) + \ln(2)\right] \\ &\quad + 2m + 2N(\alpha) + 2, \\ m &= \left\lceil \frac{1}{c}\ln\left(\frac{2}{\epsilon}\right)\right\rceil. \end{aligned}$$

Since $\Psi_k \in \overline{\Pi}_k$, the theorem is now a consequence of (6.6). ∎

7. Conclusion. We have studied KKT systems

$$\mathcal{A}_\alpha p = b$$

arising in connection with inverse problems. The spectral condition number $\kappa(\mathcal{A}_\alpha)$ of $\mathcal{A}_\alpha$ increases rapidly as the regularization parameter $\alpha > 0$ decreases, i.e., $\kappa(\mathcal{A}_\alpha) = O(\alpha^{-1})$. Our main result shows that the spectrum of $\mathcal{A}_\alpha$ is contained in three bounded

intervals with only a very limited number $N(\alpha)$ of isolated eigenvalues:

$$\mathrm{sp}(\mathcal{A}_\alpha) \subset [-b,-a] \cup [c\alpha, d\alpha] \cup \{\lambda_1, \lambda_2, \ldots, \lambda_{N(\alpha)}\} \cup [a,b],$$

provided that Tikhonov regularization is used. For severely ill posed problems $N(\alpha) = O(\ln(\alpha^{-1}))$. Due to this property of the spectrum, the MINRES method solves such KKT systems efficiently also for small values of $\alpha > 0$.

Techniques based on Chebyshev polynomials were applied in section 6 to prove that the MINRES method requires at most

$$O([\ln(\alpha^{-1})]^2) \tag{7.1}$$

iterations as $\alpha \to 0$, provided that the underlying inverse problem is severely ill posed. Our theoretical investigation concerned infinite-dimensional KKT systems with countable spectra.

The numerical examples addressed problems with state equations given in terms of discretized elliptic PDEs, which involve a mesh parameter h. Standard algebraic multigrid was used to precondition the KKT system. The MINRES method solved all the model problems studied in this paper efficiently, and the approach is robust with respect to h. Furthermore, in the numerical experiments we observed iteration counts that are almost of order

$$O(\ln(\alpha^{-1})), \tag{7.2}$$

which is better than predicted by the theoretical result (7.1). In section 6 we also discussed how the analysis leading to (7.1) can be used to explain why (7.2) is plausible in practice; one may regard (7.1) to be the asymptotic behavior as $\alpha \to 0$.

The method for solving PDE constrained optimization problems presented in this paper has so far only been tested on elliptic control problems, posed on simple geometries, subject to Tikhonov regularization and noise-free data. It must be investigated whether the technique also can handle optimization tasks involving time-dependent PDEs, inequality constraints, nonlinear problems, and other regularization schemes.

In this study we used multigrid algorithms to define suitable and rather simple preconditioners for PDE constrained optimization problems. Alternatively, one can employ more advanced preconditioners or multigrid schemes directly to the KKT system and explore whether similar nice convergence properties are valid. This would require a completely new type of investigation, but it seems plausible that a local Fourier analysis could work; see [22] and the references therein.

Appendix A. Boundedness and Babuška–Brezzi conditions. Our goal is to derive upper bounds for $\|\mathcal{A}_\alpha\|$ and $\|\mathcal{A}_\alpha^{-1}\|$. To this end, let us introduce the notation

$$\begin{aligned}
&X = H_1 \times H_2,\ \|x\| = \|(x_1,x_2)\| = \sqrt{\|x_1\|^2 + \|x_2\|^2},\\
&Y = H_2,\\
&M_\alpha = \begin{bmatrix} \alpha I & 0 \\ 0 & T^*T \end{bmatrix} : X \to X,\\
&N = \begin{bmatrix} B & A \end{bmatrix} : X \to Y,\\
&f = \begin{bmatrix} 0 \\ T^*d, \end{bmatrix}.
\end{aligned}$$

Then we can write (3.1) in the following form: Find $x = (v, u) \in X$ and $y = w \in Y$ such that

$$\begin{aligned} M_\alpha x + N^* y &= f, \\ Nx &= 0. \end{aligned}$$

This is a saddle point problem, and we can use standard techniques to analyze it; see, e.g., [10].

Note that, for any $x = (x_1, x_2) \in X$, $z = (z_1, z_2) \in X$, and $\alpha \in [0, 1]$,

$$\begin{aligned} |(M_\alpha x, z)| &\le \alpha |(x_1, z_1)| + |(T^*Tx_2, z_2)| \\ &= \alpha |(x_1, z_1)| + |(Tx_2, Tz_2)| \\ &\le \|x_1\|\|z_1\| + \|Tx_2\|\|Tz_2\| \\ &\le \|x_1\|\|z_1\| + \|T\|^2\|x_2\|\|z_2\| \\ &\le \|x\|\|z\| + \|T\|^2\|x\|\|z\| \\ &= \left(1 + \|T\|^2\right)\|x\|\|z\|. \end{aligned}$$

Also,

$$\begin{aligned} |(Nx, y)| &\le |(Bx_1, y)| + |(Ax_2, y)| \\ &\le \|B\|\|x_1\|\|y\| + \|A\|\|x_2\|\|y\| \\ &\le \|B\|\|x\|\|y\| + \|A\|\|x\|\|y\| \\ &= (\|B\| + \|A\|)\|x\|\|y\| \end{aligned}$$

for any $x = (x_1, x_2) \in X$ and $y \in Y$. Both M_α and N are thus bounded, and we conclude that

$$\|\mathcal{A}_\alpha\| \le C \quad \text{for all } \alpha \in [0, 1].$$

The coercivity of M_α on the kernel of N involves the size of the regularization parameter α. More specifically, if $z = (z_1, z_2) \in X = H_1 \times H_2$ is such that

$$Nz = 0,$$

i.e.,

$$Az_2 = -Bz_1,$$

then (2.2) implies that

$$\|z_2\| \le \tilde{C}\|z_1\|.$$

Consequently,

$$\begin{aligned} (M_\alpha z, z) &= \alpha(z_1, z_1) + (T^*Tz_2, z_2) \\ &= \alpha\|z_1\|^2 + (Tz_2, Tz_2) \\ &\ge \alpha\|z_1\|^2 \\ &\ge 0.5\alpha\|z_1\|^2 + 0.5\frac{1}{\tilde{C}^2}\alpha\|z_2\|^2 \\ &\ge \alpha\tilde{c}\|z\|^2. \end{aligned}$$

The inequalities presented in this appendix, assumption $\mathcal{A}5$, and standard theory for saddle point problems [10] assert that $\mathcal{A}_\alpha$ is continuously invertible and that

$$\|\mathcal{A}_\alpha^{-1}\| \leq \frac{1}{c\alpha} \quad \text{for all } \alpha \in (0,1].$$

Appendix B. Adjoint of the inverse Riesz map. Recall that $R_2 : H_2 \to H_2'$ is the Riesz map of the state space H_2 and that

$$\begin{aligned} &R_2^{-1} : H_2' \to H_2, \\ &(R_2^{-1})^* : H_2 \to H_2'. \end{aligned}$$

We want to show that $(R_2^{-1})^* = R_2$. To this end, let x, $y \in H_2$ be arbitrary. The associated members of the dual space are

$$x' = R_2 x \quad \text{and} \quad y' = R_2 y.$$

Then $(x', y')_{H_2'} = (x, y)_{H_2}$ and we find that

$$\begin{aligned} ((R_2^{-1})^* x, y')_{H_2'} &= (x, R_2^{-1} y')_{H_2} \\ &= (x, y)_{H_2} \\ &= (x', y')_{H_2'} \\ &= (R_2 x, y')_{H_2'}. \end{aligned}$$

Since this holds for all x, $y \in H_2$, we conclude that $(R_2^{-1})^* = R_2$.

Appendix C. Adjoint and dual operators. Let $\widehat{A} : H_2 \to H_2'$ be the operator defined in (5.5). From the diagram

$$\begin{array}{ccc} H_2 & \xrightarrow{(\widehat{A})'} & H_2' \\ \big\downarrow R_2 & & \big\uparrow R_2 \\ H_2' & \xrightarrow{(\widehat{A})^*} & H_2 \end{array}$$

we see that the relationship between the dual $(\widehat{A})'$ and adjoint $(\widehat{A})^*$ operators of $\widehat{A}$ is

$$(\widehat{A})' = R_2 \, (\widehat{A})^* R_2,$$

provided that R_2 is the Riesz map of H_2. Similarly,

$$(\widehat{B})' = R_1 \, (\widehat{B})^* R_2,$$

where R_1 is the Riesz map of H_1 and $\widehat{B} : H_1 \to H_2'$ is defined in (5.6).

Acknowledgments. It is a pleasure to thank Prof. Herbert Egger for his excellent talk at the Applied Inverse Problems conference in Vienna in 2009, which inspired the work presented in this paper. We are also grateful to the referees for excellent comments that significantly improved this paper.

REFERENCES

[1] S. S. Adavani and G. Biros, *Multigrid algorithms for inverse problems with linear parabolic PDE constraints*, SIAM J. Sci. Comput., 31 (2008), pp. 369–397.

[2] D. N. Arnold, R. S. Falk, and R. Winther, *Preconditioning discrete approximations of the Reissner-Mindlin plate model*, Math. Model. Numer. Anal., 31 (1997), pp. 517–557.

[3] O. Axelsson, *Solution of linear systems of equations: Iterative methods*, in Sparse Matrix Techniques, Copenhagen 1976, A. Dold and B. Eckmann, eds., Lecture Notes in Math. 572, Springer, New York, 1977, pp. 1–51.

[4] O. Axelsson, *Iterative Solution Methods*, Cambridge University Press, Cambridge, UK, 1994.

[5] O. Axelsson and G. Lindskog, *On the rate of convergence of the preconditioned conjugate gradient method*, Numer. Math., 48 (1986), pp. 499–523.

[6] M. Benzi, G. H. Golub, and J. Liesen, *Numerical solution of saddle point problems*, Acta Numer., 14 (2005), pp. 1–137.

[7] A. Borzì, K. Kunisch, and D. Y. Kwak, *Accuracy and convergence properties of the finite difference multigrid solution of an optimal control optimality system*, SIAM J. Control Optim., 41 (2003), pp. 1477–1497.

[8] A. Borzì and V. Schulz, *Multigrid methods for PDE optimization*, SIAM Rev., 51 (2009), pp. 361–395.

[9] A. Borzì and V. Schulz, *Computational Optimization of Systems Governed by Partial Differential Equations*, SIAM, Philadelphia, 2012.

[10] D. Braess, *Finite Elements. Theory, Fast Solvers, and Applications in Solid Mechanics*, 2nd ed., Cambridge University Press, Cambridge, UK, 2001.

[11] H. S. Dollar, N. I. M. Gould, M. Stoll, and A. J. Wathen, *Preconditioning saddle-point systems with applications in optimization*, SIAM J. Sci. Comput., 32 (2010), pp. 249–270.

[12] H. Elman, D. Silvester, and A. Wathen, *Finite Elements and Fast Iterative Solvers: With Applications in Incompressible Fluid Dynamics*, Oxford University Press, New York, 2005.

[13] H. C. Elman, *Preconditioners for saddle point problems arising in computational fluid dynamics*, Appl. Numer. Math., 43 (2002), pp. 75–89.

[14] H. C. Elman, D. J. Silvester, and A. Wathen, *Block preconditioners for the discrete incompressible Navier-Stokes equations*, Internat. J. Numer. Methods Fluids, 40 (2002), pp. 333–344.

[15] O. Ghattas, *PDE-constrained optimization at the* 2005 *SIAM conferences on CS&E and optimization*, SIAM News, 38(6) (2005), p. 8.

[16] G. H. Golub and C. F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, MD, 1996.

[17] P. C. Hansen, *Discrete Inverse Problems: Insight and Algorithms*, SIAM, Philadelphia, 2010.

[18] E. Haug and R. Winther, *A domain embedding preconditioner for the Lagrange multiplier system*, Math. Comp., 69 (1999), pp. 65–82.

[19] M. Heinkenschloss and H. Nguyen, *Neumann–Neumann domain decomposition preconditioners for linear-quadratic elliptic optimal control problems*, SIAM J. Sci. Comput., 28 (2006), pp. 1001–1028.

[20] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich, *Optimization with PDE Constraints*, Springer, New York, 2009.

[21] B. Hofmann and S. Kindermann, *On the degree of ill-posedness for linear problems with non-compact operators*, Methods Appl. Anal., 17 (2010), pp. 445–461.

[22] S. P. MacLachlan and C. W. Oosterlee, *A local Fourier analysis framework for finite-element discretizations of systems of PDEs*, Numer. Linear Algebra Appl., 18 (2011), pp. 751–774.

[23] K. A. Mardal and J. B. Haga, *Block preconditioning of systems of PDEs*, in Automated Solution of Differential Equations by the Finite Element Method, A. Logg, K. A. Mardal, and G. Wells, eds., Springer, New York, 2012, pp. 643–654.

[24] K. A. Mardal and R. Winther, *Preconditioning discretizations of systems of partial differential equations*, Numer. Linear Algebra Appl., 18 (2011), pp. 1–40.

[25] T. P. Mathew, M. Sarkis, and C. E. Schaerer, *Analysis of block matrix preconditioners for elliptic optimal control problems*, Numer. Linear Algebra Appl., 14 (2007), pp. 257–279.

[26] *Min-max Theorem*, Wikipedia, http://en.wikipedia.org/wiki/min-max_theorem.

[27] B. F. Nielsen and K.-A. Mardal, *Efficient preconditioners for optimality systems arising in connection with inverse problems*, SIAM J. Control Optim., 48 (2010), pp. 5143–5177.

[28] C. C. Paige and M. A. Saunders, *Solution of sparse indefinite systems of linear equations*, SIAM J. Numer. Anal., 12 (1975), pp. 617–629.

[29] T. Rees, H. S. Dollar, and A. J. Wathen, *Optimal solvers for PDE-constrained optimization*, SIAM J. Sci. Comput., 32 (2010), pp. 271–298.

[30] T. Rees, M. Stoll, and A. Wathen, *All-at-once preconditioning in PDE-constrained optimization*, Kybernetika (Prague), 46 (2010), p. 341–360.
[31] T. Rusten and R. Winther, *A preconditioned iterative method for saddlepoint problems*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 887–904.
[32] J. Schöberl, R. Simon, and W. Zulehner, *A robust multigrid method for elliptic optimal control problems*, SIAM J. Numer. Anal., 49 (2011), pp. 1482–1503.
[33] J. Schöberl and W. Zulehner, *Symmetric indefinite preconditioners for saddle point problems with applications to PDE-constrained optimization problems*, SIAM J. Matrix Anal. Appl., 29 (2007), pp. 752–773.
[34] D. Silvester and A. Wathen, *Fast iterative solution of stabilised Stokes systems. Part* II: *Using general block preconditioners*, SIAM J. Numer. Anal., 31 (1994), pp. 1352–1367.
[35] R. Simon and W. Zulehner, *On Schwarz-type smoothers for saddle point problems with applications to PDE-constrained optimization problems*, Numer. Math., 111 (2009), pp. 445–468.
[36] S. Takacs and W. Zulehner, *Convergence analysis of multigrid methods with collective point smoothers for optimal control problems*, Comput. Vis. Sci., 14 (2011), pp. 131–141.
[37] H. S. Thorne, *Properties of Linear Systems in PDE-Constrained Optimization. Part* I: *Distributed Control*, Tech. Report RAL-TR-2009-017, Rutherford Appleton Laboratory, 2009.
[38] H. S. Thorne, *Properties of Linear Systems in PDE-Constrained Optimization. Part* II: *Neumann Boundary control*, Tech. Report RAL-TR-2009-018, Rutherford Appleton Laboratory, 2009.
[39] F. Tröltzsch, *Optimal Control of Partial Differential Equations: Theory, Methods and Applications*, Grad. Stud. Math. 112, AMS, Providence, RI, 2010.
[40] A. Wathen, *Preconditioning for PDE-constrained optimization*, SIAM News, 43(2) (2010), p. 8.
[41] A. Weinstein and W. Stenger, *Methods of Intermediate Problems for Eigenvalues*, Academic Press, New York, 1972.
[42] G. Wittum, *Multigrid methods for Stokes and Navier-Stokes equations*, Numer. Math., 54 (1989), pp. 543–564.
[43] W. Zulehner, *Nonstandard norms and robust estimates for saddle point problems*, SIAM J. Matrix Anal. Appl., 32 (2011), pp. 536–560.